



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|-----------------|-------------|----------------------|---------------------|------------------|
|-----------------|-------------|----------------------|---------------------|------------------|

10/613,631

07/03/2003

Kuansan Wang

M61.12-0521

4356

27366

7590

09/23/2009

WESTMAN CHAMPLIN (MICROSOFT CORPORATION)

SUITE 1400

900 SECOND AVENUE SOUTH

MINNEAPOLIS, MN 55402

EXAMINER

YEN, ERIC L

ART UNIT

PAPER NUMBER

2626

MAIL DATE

DELIVERY MODE

09/23/2009

PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

|                              |                                      |                                      |  |
|------------------------------|--------------------------------------|--------------------------------------|--|
| <b>Office Action Summary</b> | <b>Application No.</b><br>10/613,631 | <b>Applicant(s)</b><br>WANG, KUANSAN |  |
|                              | <b>Examiner</b><br>ERIC YEN          | <b>Art Unit</b><br>2626              |  |

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 2 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☒ Responsive to communication(s) filed on 08 July 2009.
- 2a) ☐ This action is **FINAL**.                      2b) ☒ This action is non-final.
- 3) ☒ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☒ Claim(s) 1,6-9,11-14,16-29,33,34 and 36-38 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☐ Claim(s) \_\_\_\_\_ is/are rejected.
- 7) ☒ Claim(s) 1, 6-9, 11-14, 16-29, 33-34, and 36-38 is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some \*    c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- |  |   |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)            | 4) <input type="checkbox"/> Interview Summary (PTO-413)           |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)   | Paper No(s)/Mail Date. _____                                      |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date _____  | 6) <input type="checkbox"/> Other: _____                          |

## **DETAILED ACTION**

### ***Response to Amendment***

1. In response to the Advisory Action mailed 6/26/09, applicant has submitted an amendment and Request for Continued Examination filed 7/8/09.

Claims 1, 19, 27, and 33, have been amended. Claims 5, 10, 15, 32, and 35, have been cancelled.

### ***Response to Arguments***

To restate what was previously addressed in past Office Actions, the computer readable storage medium claims are statutory under 35 USC 101 (Amendment to the claims to change the preambles from "computer readable medium" to "computer readable storage medium" were filed 6/17/08).

In the Specification, "computer readable medium" is described as being either "computer storage media" or "communication media" (page 14, lines 10-25) where "communication media" includes signals and carrier waves (pages 14-15, paragraph starting on page 14, line 26). Computer readable media that are embodied as carrier waves and signals are non-statutory, however the computer storage media are statutory because they are limited to tangible storage devices such as RAM/ROM/etc. (page 14, lines 18-25).

Therefore, by amending "computer readable medium" to "computer readable storage medium" the non-statutory "communication media" signal/carrier embodiments

are excluded from the claim scope, and so the “computer readable storage medium” claims are statutory under 35 USC 101.

### ***Claim Objections***

1. Claims 1, 19, and 27 are objected to because of the following informalities:

Claim 1 recites “wherein the speech recognition result initialized by the SALT module” (10<sup>th</sup> to last line) but the recognition result is obtained and, by itself, has no actual processing function. It is fairly clear that applicant meant to recite speech recognition event because the event is what the SALT module initialized and so “speech recognition result” should be corrected to recite --speech recognition event—in the 10<sup>th</sup> to last line of Claim 1.

Claim 27 includes numerous intended use limitations “e.g., wherein the at least one object oriented operation initializes a recognition event to associate the speech portion of the user input with the first field and the DTMF portion of the user input with the second field” which raises a question of whether the associating “intended use” (as claimed) and whether it is actually a method step in the claim. It is fairly clear that applicant is trying to claim the association as a feature but the recognition event could be an operation to set up the association without requiring the association itself to be performed. Therefore, the claim should be clarified to remove any potential ambiguity as to whether the associating is actually a part of the claim scope. Applicant claims that the fields are “to be filled” which is presumably done by the associating of the speech input with one field and the DTMF input with the other field, and so the associating

Art Unit: 2626

limitation is not an actual method step in the claim the purpose of declaring the fields is defeated. Therefore, it is logical that the associating operation is part of the claim, but for the time being, it is only an intended use based on the claim language.

Claims 1 and 19 have similar issues.

Claim 1 recites "associating... to complete the first field" but this claim language does not necessarily require that the fields are actually filled. The association could be an intermediate step. Also, "initializ[ing] a speech recognition event having a plurality of grammars to obtain a recognition result" does not necessarily obtain the result even though it may be implied (Claim 1, lines 17-24, amended claim language)

Claim 19 recites "interrupt[ing] the form interpretation... to initialize the speech recognition event" but does not necessarily state that a speech input, used to fill the first and second VoiceXML fields, is actually obtained even if it may be implied by the intended use (8<sup>th</sup> -10th to last lines of Claim 19). Also applicant claims that invoking an object of the SALT module is for initializing the speech recognition event but does not claim that the recognition event is actually initialized.

Therefore, the intended use limitations should be clarified to ensure that there is no ambiguity that the amended limitations are actually part of the claim scope and not just descriptive of the steps clearly defined as part of the claims.

2. Appropriate correction is required.

***Allowable Subject Matter***

1. Claims 1, 6-9, 11-14, 16-29, 33-34, and 36-38, contain allowable subject matter.
2. The following is a statement of reasons for the indication of allowable subject matter:

The prior art of record generally teaches the different functions independent claims 1, 19, and 27, but does not teach the specific distribution of the functions between the SALT module and VoiceXML module and the interaction between the SALT module and VoiceXML module.

Williams et al. (US 2003/0212561) teaches programming IVR systems using both VXML/VoiceXML and SALT (paragraph 15).

As stated by applicant, Williams only mentions that VoiceXML and SALT are programming languages used in IVR systems (Amendment, page 10)

As per Claim 1, since Williams only generally teaches where VoiceXML and SALT are used in an IVR system ("computer to process information"), Williams does not teach or reasonably suggest that the VoiceXML module declares a first field and a second field and where the SALT module obtains a recognition result from an initialized recognition event with a plurality of associated grammars and associates a first portion of a recognition RESULT with a first grammar of the plurality of grammars to complete the first field declared by the VoiceXML module and associates a second portion of the recognition RESULT with a second grammar of the plurality of grammars to complete the second field.

Williams teaches using VoiceXML and SALT to implement a dialog with a corresponding call flow ("branching voice queries... caller responds with button pushes... or voice responses", paragraph 5; "call flow", paragraphs 10-11; "call flow", paragraphs 73-76; where dialog systems usually involve a window for listening to the user's response after a particular prompt is played to the user). Dialogs conducted with users necessarily include obtaining a recognition result from a speech recognition process (event) in order to determine exactly what the user is saying and to process the input properly. A dialog and call flow also necessarily has a sequence of prompts that the machine/IVR uses to communicate with the user. Since the prompts are delivered in a sequence, there is something in the VXML/SALT information that determines the order of outputting the prompts, in addition to telling the system to recognize speech and perform other functions.

Therefore, Williams teaches/suggests, by teaching a VXML/SALT IVR system, a VoiceXML module executing form interpretation and establishing an interactive dialogue with a user including instructions associated with dialog events including recognition/prompting/messaging which are executed in a defined order and since the system decides which prompt to present after a particular event without interference from a user, it automatically advances from one instruction to another instruction in a defined order. Williams also teaches/suggests a SALT module including temporal triggers. Applicant defines in the Specification that a "temporal trigger" "may include various events such as an error, exception, receipt of a message, recognition and/or no recognition or combinations thereof" (page 6, lines 13-25) and "may be triggered using

Art Unit: 2626

a listen tag that includes one or more grammar elements" (page 24, lines 7-28).

Therefore, by teaching a dialogue with a call flow which includes prompts being output to the user to obtain user speech, and recognizing user speech (e.g., Williams, Figure 7), Williams teaches temporal triggers that initialize speech recognition events that obtain recognition results. Williams suggests the portions of the claimed SALT module because, in an IVR system that uses VXML and/or SALT, the designer of the markup language document could opt to encode a portion of the functions using SALT and another portion using VXML.

Williams, however, fails to teach declaring first and second fields and distributing the portions of the recognition result to the first and second fields by associating the portions of the recognition RESULT with the GRAMMARS belonging to the first and second fields (i.e., a first grammar for the first field and a second grammar for a second field).

Aust et al. (US 5,860,059) teaches associating portions of a recognition result with a corresponding field (Figures 2A-2P; col. 3, lines 26-54; e.g., the system parses the user's answer to the question "from where to where do you want to travel" to fill a date field and a departure location [two distinct fields] with the different parts of "today from Aachen").

Aust, however, does not teach where matching the recognition result with the field is done by associating the result with a grammar of the field. In Aust, the recognition result is directly associated with the field without involving the grammars, since the grammars were already used to obtain the recognition result (i.e., the speech



Art Unit: 2626

is associated with the grammar to produce the result but the result is not associated with a grammar for a field).

Gong et al. (US 2004/0006474) teaches (Figures 17, 21, and 25, paragraph 245; paragraph 251) describes a VXML interface including a city grammar, a state grammar and a street grammar (paragraph 251). Since each of the city/state/street grammars can include Washington (Washington St./Washington DC/Washington state), this suggests that the system should associate an input of Washington with a particular grammar to determine if it is a city, a state, or a street.

However, Gong resolves this by making only one grammar active at a given time (paragraphs 255-256). Since the recognition result is obtained from a particular grammar, there is no need to associate the result with the grammar.

Even associating the recognition result with the grammar is obvious, the prior art of record does not specifically teach that the SALT portion of a markup language document, particularly, accomplishes these steps. Williams suggests that the feature could be implemented in SALT but there is no apparent reason for one of ordinary skill in the art to do so without employing impermissible hindsight (Amendment, page 10)

Therefore, the prior art of record does not teach associating portions of the recognition result with a particular grammar to complete one of the fields declared by the VoiceXML module, where this association, along with the other claimed functions of the SALT module, is performed by the SALT module, in combination with the remaining limitations in Claim 1 (including the assigned tasks of the VoiceXML module).

As per Claim 19, Williams and Aust teach/suggest performing speech recognition events, prompting events, etc., in a dialog to fill declared fields using a VXML/SALT IVR system, where voice markup language documents are processed in the order that the instructions in the document are written in to produce the dialog, which includes recognition and prompting, as discussed above regarding Claim 1. The inherent features of processing voice markup language documents involve automatically advancing/moving through an ordered list of instructions and performing functions based on the markup language tags and other command strings in the voice markup language document, which are claimed in Claim 1.

Aust further suggests looping through the VoiceXML executable instructions in a defined order until the first and second VoiceXML field have been filled by the user because Aust teaches continuously prompting for missing information depending on what the user says and what the system still needs to know to complete a transaction (i.e., ask for the missing destination, time, etc.) (Figures 2A-2P; col. 3, lines 26-54).

Taylor (US 6,922,411) teaches looping through VXML instructions (col. 16-17. Table 6A; especially col. 17, "a looping structure so that vxml elements can be repeated ...before timing out") in order to ensure input (i.e., ensure that a field is filled) where the field in Taylor is whatever memory location is used to contain the user's speech. The loop also for controlling prompting events because Taylor teaches playing an audio prompt X times.

Loops and interrupts in programmed source codes are also well-known in the art. It is also well-known that markup language parsers read through markup language

Art Unit: 2626

documents in some form of sequence and encounter markup language tags and perform a function associated with an encountered tag (e.g., performing some sort of display in response to detecting an <HTML> tag). The interrupting of a loop through a markup language document and automatically advancing/moving to a subsequent instruction in a defined order is not new and/or is obvious.

The prior art of record, however, does not teach/suggest that the SALT module specifically handles the speech recognition events while the VoiceXML module handles the prompting events and declares VoiceXML fields (instead of, for example, having the SALT tags declare fields or using some other markup language to declare fields). Even though one of ordinary skill in the art could design a voice markup language document using a combination of SALT and VoiceXML/VXML (where Williams teaches combining SALT and VoiceXML in a voice markup language system), there is no apparent reason for one of ordinary skill in the art to require the SALT and VoiceXML modules to perform their respective functions as defined in Claim 19 without employing impermissible hindsight (Amendment, page 10). Applicant's claim 19 is a species and/or an element of the broader genus of SALT/VXML documents described in Williams because it specifically defines the functions performed by the SALT module and the VXML module.

As per Claim 27, Williams and Aust, as discussed above, teach/suggest the use of voice markup language documents with SALT and VoiceXML, which contain sequentially ordered operations/instructions for conducting a dialog (including

Art Unit: 2626

recognition and prompting), and also where the dialog fills in fields corresponding to inputs.

Zou et al. (US 6,246,983) further teaches multi-modal inputs using DTMF and voice (col. 3, line 51 – col. 4, line 3).

Chaves (US 6,510,414) teaches where the grammar of the recognition system can recognize both DTMF signals and speech (col. 4, line 65 - col. 5, line 10) which implies that an input can include both DTMF signals and speech

Chang et al. (US 2003/0149565) teaches parsing an input which can include DTMF signals and/or spoken sounds (paragraph 9) in the context of a location system ("Mapquest server", paragraph 155) which implies a set of fields such as ZIP codes, street names, city names, etc. (paragraphs 153-154). This suggests dividing/parsing a DTMF/Speech combination input including a DTMF ZIP code and spoken street/city name to fill a city/street and ZIP code field because there is less risk of erroneous recognition to enter a ZIP code manually and it is much more tedious to enter a city/street name by hand as opposed to speaking it. It also logically follows that the system would fill a ZIP code field with a DTMF ZIP code and the city/street name field with a spoken city/street name because to do anything else would constitute an error (i.e., the different components of the input are put where they are supposed to be).

Brotman et al. (US 2001/0049599) teaches where ZIP codes and other numerical data are entered using DTMF codes (paragraph 10).

However, similar to the reasoning discussed above regarding claims 1 and 19, the prior art of record does not teach or reasonably suggest where associating a spoken

Art Unit: 2626

portion of an input with one field, and where a DTMF portion of an input is associated with a second field, is performed by a SALT module, where the fields are declared by the VoiceXML module, and where the SALT and VoiceXML modules perform the other functions assigned to them in Claim 27.

Even though one of ordinary skill in the art could design a voice markup language document using a combination of SALT and VoiceXML/VXML (where Williams teaches combining SALT and VoiceXML in a voice markup language system), there is no apparent reason for one of ordinary skill in the art to do so without employing impermissible hindsight (Amendment, page 10)

In summary, the actual method steps of claims 1, 19, and 27, are not new or non-obvious but the prior art of record does not teach or suggest the combinations where these method steps are distributed between the SALT module and VoiceXML module in the manner defined in Claims 1, 19, and 27.

### ***Conclusion***

3. This application is in condition for allowance except for the following formal matters:

Objections to claims 1, 19, and 27.

Prosecution on the merits is closed in accordance with the practice under *Ex parte Quayle*, 25 USPQ 74, 453 O.G. 213, (Comm'r Pat. 1935).

Art Unit: 2626

A shortened statutory period for reply to this action is set to expire **TWO MONTHS** from the mailing date of this letter.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to ERIC YEN whose telephone number is (571)272-4249.

The examiner can normally be reached on M-F 7:30-4:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on 571-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

EY 9/21/09

/Richemond Dorvil/

Supervisory Patent Examiner, Art Unit 2626